



**Finansuoja
Europos Sąjunga**
NextGenerationEU

DI PRODUKTO SUKŪRIMAS I STADIJA ATASKAITA

Projekto pavadinimas: *Dirbtinio intelekto technologijų įgalinta rinkos ir konkurencinės aplinkos stebėsenos ir analitikos platforma*

Projekto numeris: *02-018-K-0121*

Projekto vykdytojas: *UAB „Porta futuri“*

Ataskaitos data: *2025-03-19*

Lapų skaičius: *30*



**NAUJOS KARTOS
LIETUVA**

1. Veiklos tikslas	4
2. Veiklos užduotys	4
3. Rinkos poreikis ir projekto svarba	4
3.1. Rinkos analizės rezultatai	4
3.2. Projekto svarba	7
4. Technologiniai neapibrėžtumai ir iššūkiai	7
4.1. DI „haliucinacijos“ problema	7
4.2. Ribota produkto informacija	7
4.3. Integracijos su esamomis sistemomis iššūkiai	8
4.4. „Human in the loop“	9
5. Tyrimų metodika	9
5.1. Rinkos poreikių tyrimas	9
5.2. Kokybiniai interviu	9
5.3. Duomenų analizės ir paruošimo tyrimas	10
5.4. Konceptualus vartotojo sąsajos maketavimas	10
6. Tyrimo eiga	10
6.1. Pirminiai kokybiniai interviu ir duomenų analizė	10
6.1.1. Kokybiniai interviu (4 pokalbiai)	10
6.1.2. Prekių katalogų duomenų analizė	11
6.1.3. Rezultatas	11
6.2. Rinkos poreikių tyrimas ir konkurentų sprendimų analizė	11
6.2.1. Mokslinės literatūros ir statistinių šaltinių analizė	11
6.2.2. Konkurentų ir rinkoje egzistuojančių sprendimų apžvalga	12
6.2.3. Tendencijų ir pagrindinių iššūkių identifikavimas	12
6.2.4. Rezultatai	12
6.3. DI technologinių variantų analizė	13
6.3.1. Privalumų ir trūkumų įvertinimas	13
6.3.2. Eksperimentinis bandymas	14
6.3.3. Sprendimas dėl maketo	14
6.4. Antrieji kokybiniai interviu su kitais ekspertais	14
6.4.1. Rezultatai ir pagrindinės išvados	15
6.4.2. Poveikis projekto kryptčiai	15
6.5. Vartotojo sąsajos maketavimas ir testavimas	15
6.5.1. Maketo kūrimo eiga	16
6.5.1.1. Žemo detalumo (low-fidelity) maketai	16
6.5.1.2. Aukšto detalumo (high-fidelity) maketai	16
6.5.1.3. Human in the loop	19
6.5.2. Maketo testavimas su vidiniais įmonės darbuotojais	20
6.5.2.1. Testinių scenarijų pavyzdžiai	20
6.5.2.2. Surinkti rezultatai	21
6.5.2.3. Korekcijos	22
6.3. Maketui pritaikyta architektūra	23
6.4. Tyrimo rezultatų išvados ir tolesni veiksmai	24
7. Pasiekti rezultatai	25
7.1. Sukurtas vartotojo sąsajos maketas	25
7.2. Patvirtintas „human in the loop“ naudingumas	25

7.3. Gauti pirmieji testų rodikliai	26
7.4. Sukurtas tolesnis plėtros planas	26
7.5. Verslo naudą patvirtinančios įžvalgos	27
8. Veiklos vykdymo metu atliktų tyrimų eigos pakeitimai	28
8.1. Modelio pasirinkimo korekcija	28
8.2. Metodologijos pritaikymas esant duomenų trūkumui	28
8.3. Skirtingų modelio versijų palyginimas	29
8.4. Testavimo apimties korekcija	29
9. Kita su projekto įgyvendinimu susijusi informacija (ekranvaizdžiai, struktūrinės schemos ir pan.)	29
9.1. Gerosios praktikos analizė ir analogų įžvalgos	29
9.2. Sistemos architektūros vizualizacijos ir techninė konfigūracija	30
9.3. Naudoti įrankiai ir technologijos	30
9.4. Projekto įgyvendinimo dinamika ir Agile adaptacijos	30

- **1. Veiklos tikslas**

Atlikti išsamią rinkos poreikių analizę, ištestuoti maketą, siekiant patvirtinti technologinės koncepcijos įgyvendinamumą ir sisteminį sprendimo sudedamųjų dalių veikimą.

- **2. Veiklos užduotys**

Įgyvendinant personalizuoto DI paieškos maketo kūrimą, buvo išskirti keli pagrindiniai veiksmi:

- **Rinkos ir vartotojų analizė:** ištirti esamų e. komercijos paieškos sprendimų trūkumus, vartotojų nepasitenkinimo priežastis ir personalizavimo galimybes (remtasi viešai prieinamais tyrimais ir statistika apie paieškos naudojimą ir efektyvumą). Ši užduotis apima literatūros ir konkurentų analizę, siekiant pagrįsti sprendimo svarbą, naudojimo atvejų (use case) identifikavimą, atlikti kokybinius interviu su potencialiais klientais, kad aiškiai suformuluotume būsimos sistemos funkcijas atliepančias rinkos poreikius.
- **Sprendimo koncepcijos parengimas:** apibrėžti personalizuotos paieškos modelio viziją, numatanti, kaip DI galėtų pritaikyti paieškos rezultatus individualiam vartotojui (pvz., pagal jo naršymo istoriją, preferencijas). Taip pat, suformuluoti, kaip DI sprendimas integruotųsi su esamomis e.komercijos sistemomis.
- **Maketo kūrimas ir realizavimas:** sukurti pradinį DI maketą, naudojant didelio kalbos modelio technologiją bei leidžiantį vartotojams sąveikauti su paieškos ir rekomendacijų sistema natūralia kalba.
- **Testavimas ir derinimas:** išbandyti maketą vidinėje aplinkoje bei su galutiniais vartotojais (jei įmanoma) siekiant įvertinti jo tikslumą, naudotojų patirtį ir verslo vertę. Rinkti metrikų duomenis (paieškos tikslumą, atsakymo greitį, paspaudimų ir konversijų rodiklius) bei gauti grįžtamąjį ryšį, siekiant identifikuoti pagrindinius sistemos tobulinimo aspektus.
- **Rezultatų analizė:** įvertinti surinktus kiekybinius bei kokybinius duomenis, siekiant nustatyti DI sprendimo efektyvumą ir parengti rekomendacijas tolimesniam maketo tobulinimui, įskaitant papildomų taikomųjų mokslinių tyrimų poreikius prieš pilną integraciją į realią verslo aplinką.

- **3. Rinkos poreikis ir projekto svarba**

3.1. Rinkos analizės rezultatai

Elektroninės komercijos rinka sparčiai auga - 2023 m. pasauliniai e. prekybos pardavimai siekė apie 6,5 trilijono USD, o iki 2027 m. prognozuojama, kad ši vertė išaugs ~47% ir priartės prie 10 trilijonų USD ([redstagfulfillment.com](https://www.redstagfulfillment.com)). Šis augimas reiškia milžinišką prekių įvairovę, su kuria susiduria pirkėjai. Vartotojams tampa sunku atrasti jiems aktualiausias

prekes: tyrimai rodo, kad daugiau nei pusė (54%) pirkėjų palieka internetinę parduotuvę, jei prekių pasirinkimas atrodo per daug sudėtingas, o 42% net atsisako planuoto pirkinio dėl „per didelio pasirinkimo“ sukeltos sprendimų paralizės (globo.newswire.com). Pirkėjai tikisi personalizuotos pagalbos - net 68% apklaustųjų nori, jog jų mėgstami pardavėjai suteiktų asmeninių patarimų ir produktų rekomendacijų internetu (globo.newswire.com).

3.1.1. Verslo problematika

Auganti rinka atskleidžia, kad daugelis elektroninės komercijos įmonių susiduria su nepakankama produktų informacija ir netiksliais rekomendavimo algoritmais. Jei prekių aprašymai ir duomenys nėra išsamūs, algoritmai neturi pakankamai "žinių", kad galėtų tiksliai siesti produktus su vartotojų poreikiais. Dėl to pirkėjams gali būti pateikiami neaktualūs arba neįdomūs pasiūlymai, kas blogina vartotojų patirtį ir mažina pardavimus.

Be to, elektroninės komercijos atstovai pažymi, kad prekių puslapiuose pateikta informacija, jei ir yra teikiama, dažnai nėra pakankamai struktūruota. Pavyzdžiui, naudojant trečiųjų šalių įskiepius, tokius kaip Flixxmedia, kuri specializuojasi prekės ženklo stiprinančio, interaktyvaus turinio kūrimo ir gamintojų teikiamų vizualių sprendimų integravime, prekės puslapiai gali būti dinamiškai atnaujinami. Tačiau, skirtingai nuo produktų informacijos valdymo (PIM) sistemų, kurios yra vidiniai duomenų centrai, tokie įrankiai pateikia gamintojų teikiamą turinį be poreikio iš prekybos atstovų pusės ranka įvesti duomenis. Dėl to nestruktūruoti duomenys nėra pritaikomi nei prekių paieškai, nei palyginimui, nei kitiems prekių duomenų gavimo scenarijams.

Šios **verslo problemos lemia ne tik vartotojų, bet ir darbuotojų iššūkius**. Tiek klientai, tiek įmonės darbuotojai naudoja tą pačią ribotą duomenų bazę sprendžiant užklausas.

3.1.2. Personalizuota patirtis

Verslo aplinkoje itin vertinama **personalizacija** - gebėjimas kiekvienam naudotojui pateikti pritaikytus rezultatus. DI paieškos sistemos mokosi iš naudotojų elgsenos (naršymo istorijos, ankstesnių paieškų, pirkimų) ir pagal tai **pritaiko rezultatus individualiems poreikiams** (searchanise.io). Tarkime, jei pirkėjas anksčiau įsigijo tam tikro modelio telefoną, ieškant „telefono dėklas“ jam pirmiausia bus parodyti būtent to modelio dėklai (searchanise.io). Personalizuota paieška leidžia vartotojui jaustis suprantamam - jam rodomi *aktualūs* pasiūlymai, o ne bendrinė masė. Tai ne tik gerina patirtį, bet ir skatina didesnę įsitraukimą bei konversijas (searchanise.io). Gartner duomenimis, DI diegimas el. prekyboje vidutiniškai daugiau nei 25% pagerina klientų pasitenkinimą ir pajamas (bloomreach.com), daugiausia dėl personalizuotų rekomendacijų ir patogesnės sąveikos.

3.1.3. Geriau tvarkomi duomenys

DI sistemos gali automatiškai **kategorizuoti ir indeksuoti turinį**, todėl paieška veikia patikimiau. Užuo rankiniu būdu priskyrus raktinius žodžius kiekvienam straipsniui ar prekei, mašininio mokymosi algoritmai patys išmoksta, kaip turinys susijęs tarpusavyje (thevirtualforge.com). Tai mažina žmogiškų klaidų tikimybę ir užtikrina, kad vartotojams būtų parodyti visi atitinkami rezultatai. Be to, DI pašalina šališkumus, kurie galėtų atsirasti dėl žmogaus klaidingo duomenų interpretavimo - rezultatai labiau objektyvūs (thevirtualforge.com).

3.1.4. Natūralios kalbos užklausų supratimas

DI paieškos sistemos pasitelkia natūralios kalbos apdorojimą (NLP), kad geriau suprastų naudotojo ketinimą užklausoje. Tai reiškia, jog vartotojai gali ieškoti informacijos formuluodami klausimus panašiai kaip bendraujant su žmogumi, o paieškos variklis

supranta kontekstą ir pateikia atitinkamus atsakymus ([shopify.com](https://www.shopify.com)). Pavyzdžiui, užuot ieškoję fragmentiškai „žalia ilga suknelė“, vartotojai gali įvesti išsamų aprašymą („ieškau ilgos miško žalumos spalvos suknelės su V formos iškirpte ir ilgomis rankovėmis...“) - DI sistema supras tokios detalios užklauskos reikšmę ir pateiks konkrečią atitinkančią prekę ([shopify.com](https://www.shopify.com)). Toks **užklauskų interpretavimas** didina paieškos patogumą, nes vartotojui nereikia prisitaikyti prie sistemos - sistema prisitaiko prie vartotojo.

3.1.5. Pagrindiniai DI paieškos sistemų trūkumai

- **Didelės diegimo sąnaudos.** Pažangios DI paieškos įgyvendinimas reikalauja nemažų investicijų į technologijas, specialistus ir infrastruktūrą. Mažoms ir vidutinėms įmonėms tai gali būti finansiškai **sudėtinga** ([searchanise.io](https://www.searchanise.io)). Reikia ne tik sukurti ar įsigyti algoritmus, bet ir juos nuolat palaikyti, treniruoti su naujais duomenimis. Nors ilgalaikėje perspektyvoje DI gali sumažinti veiklos kaštus (pvz., automatizuodamas klientų aptarnavimą), pradinė investicija gali atsipirkti ne iš karto.
- **Privatumo klausimai.** DI sistemos efektyvumui būtini dideli duomenų kiekiai apie vartotojų elgesį, preferencijas. Tai kelia **vartotojų duomenų privatumo** problemų ([searchanise.io](https://www.searchanise.io)). Klientai gali jaustis nejaukiai, jei supranta, kad jų paieškos ir pirkimo istorija yra nuodugniai analizuojama. Verslai, diegiantys DI paiešką, turi užtikrinti duomenų saugumą ir būti skaidrūs, kaip tie duomenys naudojami, kitaip rizikuojama prarasti klientų pasitikėjimą.
- **Šališkumo ir tikslumo iššūkiai.** DI algoritmų kokybė priklauso nuo mokymosi duomenų. Jei duomenys **nepilni ar šališki**, paieškos rezultatai taip pat gali būti iškreipti ([searchanise.io](https://www.searchanise.io)). Tai gali lemti, kad tam tikroms naudotojų grupėms rodomi mažiau tinkami ar net diskriminaciniai rezultatai. Be to, net gerai apmokytos sistemos nėra neklystančios - „**haliucinacijų**“ problema, kai DI sugeneruoja įtikinamai skambančią, bet klaidingą informaciją, gali pakenkti vartotojo patirčiai. Pavyzdžiui, DI gali neteisingai suprasti produkto techninę charakteristiką ir pateikti klaidinančią detalę (kaip pastebėta *Perplexity AI* atveju, kur DI supainiojo „1,20 FT“ su „120 FT“, klaidingai suprasdamas matavimo vienetų reikšmę) ([hypotenuse.ai](https://www.hypotenuse.ai)). Todėl būtina nuolat tikrinti ir tobulinti algoritmų tikslumą.
- **Poreikis nuolat mokytis ir prižiūrėti.** DI paremta paieška nėra „įdėk ir pamiršk“ sprendimas. Kad ji veiktų gerai, reikia nuolatinio **modelių mokymo** su naujausiais duomenimis, algoritmų derinimo. Netinkamai apmokyti modeliai gali duoti labai prastus rezultatus ([searchanise.io](https://www.searchanise.io)). Personalizacijos algoritmai taip pat reikalauja laiko prisitaikyti, jei vartotojo elgsena staiga pasikeičia (pvz., pradeda ieškoti visiškai naujos produktų kategorijos), sistema ne iš karto supras naujus ketinimus ([searchanise.io](https://www.searchanise.io)). Tam tikrais atvejais tai gali trumpam pabloginti paieškos aktualumą, kol DI „persimokys“ prie naujo vartotojo profilio. Verslui tai reiškia būtinybę skirti resursų stebėti DI veikimą, koreguoti parametrus ir, jei reikia, įsikišti su žmogaus ekspertize, kad išvengtų didelių nukrypimų.
- **Vartotojų priėmimas ir pasitikėjimas.** Nors DI paieškos populiarėja, dalis naudotojų gali būti nepasiruošę visiškai pasikliauti automatizuotais atsakymais. Tyrimai rodo, kad apie pusę vartotojų vis dar dvejoja, ar pereiti prie vien DI varomų paieškos sistemų ([statista.com](https://www.statista.com)). Kai kurie gali pasigesti tradicinio paieškos rezultatų sąrašo, norėti patys peržiūrėti kelis informacijos šaltinius. **Pasitikėjimo** klausimas svarbus: jei DI generuoja atsakymą, vartotojas gali nežinoti, kokiais šaltiniais remtasi. Dėl to modernios DI paieškos (pvz., *Bing AI*, *Perplexity*) stengiasi **pateikti citatas ar nuorodas** kartu su sugeneruotais atsakymais ([shopify.com](https://www.shopify.com),

hypotenuse.ai), kad naudotojas galėtų patikrinti informacijos pagrindą. Nepaisant to, įmonėms diegiant šiuos sprendimus svarbu edukuoti vartotojus, kaip naudotis DI paieška ir ką daryti, jei gautas atsakymas kelia abejonių.

3.2. Projekto svarba

Šio projekto sprendimas - personalizuotų rekomendacijų sistema - yra svarbus atsakas į minėtus iššūkius. Tinkamai pritaikius DI, galima padėti vartotojams greičiau rasti norimas prekes ir sumažinti informacijos perkrovą, tuo pačiu didinant konversijas.

• 4. Technologiniai neapibrėžtumai ir iššūkiai

Tiriant personalizuotos paieškos modelius, susidurta su keliais techniniais neapibrėžtumais ir iššūkiais.

4.1. DI „haliucinacijos“ problema

Generatyviniai dirbtinio intelekto modeliai (ypač dideli kalbos modeliai) sukurti tam, kad pateiktų įtikinamai atrodančią informaciją - dažnai tai yra ir tikslu, o modelis teikia atsakymus, kurie atitinka realybę. Tačiau esant tam tikroms sąlygoms (pvz., jei užklaustos formuluotė yra neįprasta arba duomenų bazėje trūksta išsamios informacijos), DI modeliai gali generuoti atsakymus, kurie, nors ir atrodo įtikinami, faktiškai yra neteisingi arba neatitinka tikrų duomenų. Toks fenomenas vadinamas *haliucinacijomis* (en.wikipedia.org). Pavyzdžiui, modelis gali rekomenduoti prekę, kuri nėra įtraukta į katalogą, arba pateikti neteisingus teiginius apie prekių savybes.

Įtaka rekomendacijoms yra reikšminga. Jei DI sistema pateikia klaidinančius atsakymus, vartotojų pasitikėjimas paieškos rezultatais gali sumažėti. Tai yra ypač rizikinga, kai rekomenduojamos prekės neatitinka vartotojų poreikių ar aprašomos netinkamai, dėl ko vartotojas gali likti nusivylęs ir prarasti pasitikėjimą paieškos sistema.

Projekto metu ši problema identifikuota kaip esminis technologinis neapibrėžtumas - reikia spręsti, **kaip užtikrinti atsakymų tikslumą ir sumažinti „haliucinacijų“ tikimybę.**

4.2. Ribota produkto informacija

Personalizuotos paieškos efektyvumas glaudžiai priklauso nuo turimų duomenų kokybės ir išsamumo. Rinkos analizė ir interviu su elektroninės komercijos atstovais atskleidė, kad daugelis įmonių susiduria su problema - prekių aprašymai ir specifikacijos dažnai yra nepakankamai detalūs arba visiškai trūksta. Tai lemia, kad rekomendavimo algoritmai neturi pakankamai „žinių“, kad tiksliai siėtų produktus su vartotojų poreikiais, o rezultatas gali būti neaktualūs arba neįdomūs pasiūlymai, kurie blogina vartotojų patirtį ir mažina pardavimus.

Be to, interviu metu pabrėžta, kad prekių puslapiuose pateikta informacija, jei ir teikiama, dažnai nėra pakankamai struktūruota. Pavyzdžiui, naudojami trečiųjų šalių įskiepiai, tokie kaip Flixmedia, gali dinamiškai atnaujinti turinį, tačiau skirtingai nuo PIM sistemų, kurių duomenys yra centralizuoti, tokia informacija nėra integruota į prekių paieškos ar

palyginimo funkcionalumus. Todėl nestruktūruoti duomenys nėra pritaikomi nei paieškai, nei palyginimui, nei kitiems produktų duomenų gavimo scenarijams.

Ši ribota produkto informacija kelia dvi pagrindines problemas:

- DI modeliui **trūksta faktinės informacijos**, kuria jis galėtų tiksliai generuoti rekomendacijas.
- **Padidėja „haliucinacijų“ rizika**, nes modelis, stengdamasis užpildyti informacijos spragas, gali pateikti neteisingus teiginius apie prekių savybes ar net pasiūlyti produktus, kurie iš tikrųjų nėra kataloge.

Poveikis rekomendacijų tikslumui reikšmingas. Tyrimai rodo, kad apie 30% pirkėjų apskritai atsisako pirkti, jei neranda pakankamai informacijos apie produktą (bluemeteor.com). Todėl ribota produkto informacija ne tik mažina rekomendacijų tikslumą, bet ir reiškia prarastas pardavimų galimybes bei blogina tiek klientų, tiek darbuotojų patirtį, nes abu naudoja tą pačią ribotą duomenų bazę pateikiant užklausas.

Šio iššūkio kontekste projekte nesvarstyta, kaip pagerinti produktų duomenų bazės kokybę, nes mūsų verslo modelis neleidžia spręsti šios problemos.

4.3. Integracijos su esamomis sistemomis iššūkiai

Interviu metu klientų aptarnavimo bei telemarketingo vadovai atskleidė, kad dabartinės sistemos dažnai veikia atskirai, o tai lemia ilgesnį paieškos laiką bei informacijos fragmentaciją.

AI sistemų integracija su esama infrastruktūra gali būti sudėtinga - tai patvirtina ir industrijos patirtis (the-future-of-commerce.com). Mūsų atveju iššūkiai apima:

- **Duomenų sinchronizavimą realiu laiku:** kad rekomendacijos būtų aktualios, DI modelis turi nuolat gauti naujausią informaciją apie prekių likučius, kainas, naujus produktus. Sudėtinga užtikrinti spartų ir patikimą dvipusį ryšį tarp DI modelio ir vidinių sistemų (pvz., prekių katalogo, kliento veiksmų istorijos metrikų).
- **Veikimo greitį ir mastelį:** rekomendacijų sistemai reikia apdoroti didelius duomenų kiekius (tūkstančius vartotojo veiksmų istorijos metrikų) ir pateikti atsakymus per kelias sekundes. Reikia optimizuoti modelio užklausas, kad *pokalbio sąsajoje* vartotojas nelauktų per ilgai. Tai kelia technologinio neapibrėžtumo - ar modelis spės generuoti atsakymą pakankamai greitai, kai vienu metu sistema naudosis daug vartotojų.
- **Saugumas ir stabilumas:** bet kokia integracija turi nepažeisti duomenų saugumo. DI modeliui gavus prieigą prie tam tikrų vartotojų duomenų (pvz., ankstesnių pirkimų istorijos) turi būti užtikrinta, kad tie duomenys būtų tvarkomi laikantis BDAR (*GDPR*) ir įmonės vidaus politikos. Taip pat svarbu, kad DI komponento klaida nesugadintų

vartotojo patirties (pvz., geriau jau nerodyti jokios rekomendacijos nei rodyti klaidingą informaciją).

4.4. “Human in the loop”

Žmogiškoji kontrolė (*Human in the loop*) - tai priemonė, užtikrinanti, kad DI generuojamos rekomendacijos būtų peržiūrimos ir patvirtinamos specialistų, siekiant sumažinti ribotos produkto informacijos ir DI „haliucinacijų“ riziką. Šis metodas leidžia:

- **Rezultatų kokybės gerinimą:**

Kai DI modeliai, kaip aprašyta 4.1 ir 4.2 poskyriuose, kartais generuoja netikslius arba nepakankamai išsamius atsakymus dėl trūkstamos informacijos, patyrusių specialistų įsikišimas leistų užtikrinti, kad prieš rekomendacijų pateikimą galutiniam vartotojui, atsakymai būtų įvertinti. Tai padidintų rekomendacijų patikimumą ir mažintų klaidų skaičių.

- **Efektyvaus grįžamojo ryšio užtikrinimą:**

Žmogiška kontrolė veikia kaip papildomas kontrolės mechanizmas, kuris greitai identifikuoja netikslumus DI generuojamuose atsakymuose. Specialistų atsiliepimai padėtų koreguoti modelio parametrus ir tobulinti algoritmą, užtikrinant aukštesnį rekomendacijų tikslumą.

- **Integracijos su DI modeliais sinergiją:**

Integravus „Human in the loop“, galima pasinaudoti tiek DI modelių galimybėmis, tiek žmonių sprendimų priėmimo patirtimi. Tai ypač svarbu siekiant užtikrinti, kad net jei DI modelis pateikia įtikinamą, bet netikslią informaciją, specialistų peržiūra užkirstų kelią klaidingų rekomendacijų pateikimui. Šis metodas yra gerai žinomas *Machine Learning* (ML) praktikoje ir laikomas būtinu, siekiant nuosekliai didinti rezultatų kokybę bei užtikrinti efektyvesnį grįžtamąjį ryšį.

- **5. Tyrimų metodika**

Siekiant išsamiai ištirti DI sprendimo savybes ir įvertinti sukurtos personalizuotos paieškos sistemos maketo veikimą, buvo taikytos **keturios** metodinės priemonės.

5.1. Rinkos poreikių tyrimas

Analizuojama mokslinė literatūra bei viešai prieinami duomenys apie elektroninės komercijos paieškos naudojimo tendencijas, vartotojų elgseną ir personalizavimo poveikį. Taip pat lyginami konkurentų sprendimai, siekiant nustatyti esamas spragas ir pagrindinius kriterijus, kuriuos turėtų atitikti mūsų sprendimas.

5.2. Kokybiniai interviu

Kokybiniai interviu yra tyrimų metodika, kuri leidžia išgauti giluminę ir kontekstinę informaciją apie dalyvių patirtį, nuomones ir lūkesčius. Šio metodo esmė - struktūruotas arba pusiau struktūruotas pokalbis su ekspertais, vartotojais ar organizacijos atstovais, kuris padeda atskleisti subtilius problematikos niuansus, kai negalima remtis kiekybiniais duomenimis.

5.3. Duomenų analizės ir paruošimo tyrimas

Atliekamas vidinių (įmonės turimų duomenų) produktų duomenų auditavimas ir analizė. Tikslas - įvertinti, kiek išsamūs bei struktūruoti esami duomenys, ir koku mastu jie gali būti panaudojami DI modelio mokymui. Tyrimo metu nagrinėjami:

- Produktų katalogo duomenys (katalogas, kategorijos, aprašymai);
- Vartotojų elgsenos duomenys (pirkimų istorija, vartotojo metrikos);
- Duomenų integravimo iššūkiai.

Šis tyrimas padeda aiškiai suprasti duomenų kokybės problemas ir kaip jos gali paveikti DI modelio veikimą.

5.4. Konceptualus vartotojo sąsajos maketavimas

Šio ankstyvo kūrimo etapo metu, kai dar nevyksta intensyvus programavimas, pirmiausia sukuriama paprasti ekrano išdėstymo brėžiniai (*wireframes*). Juose vaizduojama būsimos sistemos informacijos struktūra, mygtukų vieta, rekomenduojamų produktų ir filtrų skyriai. Tokius maketus testuojant su realiais komandos dalyviais, jie gauna konkrečias užduotis, pavyzdžiui, „surasti konkretų produktą pagal tam tikrus kriterijus“ arba „palikti komentarą apie rekomenduojamą planą“. Stebima, ar naudotojams aišku, kaip naudotis siūlomomis funkcijomis, o visi surinkti pastebėjimai padeda greitai koreguoti išdėstymą. Kadangi maketai parengti tik brėžinių pavidalu, keitimus juose daryti paprasta ir nebrangu. Tai leidžia anksti identifikuoti vartotojo sąsajos silpnąsias vietas, sutaupyti išlaidų vėlesniuose etapuose ir užtikrinti, kad galutinis sprendimas geriau atitiktų naudotojų poreikius.

• 6. Tyrimo eiga

Siekdami įgyvendinti 2 *skyriuje* iškeltus uždavinius ir įvertinti kuriamos personalizuotos DI paieškos maketo veiksmingumą, taikėme visus 5 *skyriuje* minimus tyrimų metodus (rinkos poreikių tyrimą, kokybinius interviu, duomenų analizę, maketavimą). Toliau pateikiama nuosekli tyrimo eiga, parodanti, kaip sprendėme identifikuotus techninius neapibrėžtumus (4 *skyrius*) ir kokiomis priemonėmis užtikrinome, kad galutinis maketas atliktų realius rinkos bei verslo poreikius.

6.1. Pirminiai kokybiniai interviu ir duomenų analizė

Pradinis žingsnis - išsiaiškinti bendras telekomunikacijų ir elektroninės komercijos rinkos atstovų patirtis, susijusias su paieškos optimizacija bei personalizacijos svarba. Taip pat įvertinti, kiek realiai organizacijose yra galimybių naudoti DI sprendimus, kokie duomenys apie klientus ir produktus kaupiami, ar jie pakankamai struktūruoti.

6.1.1. Kokybiniai interviu (4 pokalbiai)

Pirmiausia buvo apklausti keturi ekspertai:

- Manager of Client Intelligence and Client Service (telekomunikacijų srityje),
- El. prekybos pardavimų skatinimo projektų grupės vadovas (e. komercijos srityje),
- E-commerce executive (e. komercijos srityje),

- Head of Analytics (BI & AI) (telekomunikacijų srityje).

Šių specialistų pasirinkimas pagrįstas skirtingais vaidmenimis vertės grandinėje: nuo klientų aptarnavimo iki verslo analizės ir e. komercijos vadovų. Toks požiūrių spektras leido geriau suprasti organizacinius iššūkius, bei vartotojų ir darbuotojų poreikius.

Interviu metu išsiaiškinome:

- Paieškos optimizacijos poreikį ir dabartines spragas. Daugelis patvirtino, kad esamos paieškos sistemos vartotojams dažnai pateikia neaktualius rezultatus dėl ribotų produktų aprašymų ir nelanksčių filtravimo mechanizmų.
- Personalizacijos galimybes. Buvo pabrėžta, kad personalizuotų pasiūlymų kūrimas galėtų reikšmingai prisidėti prie pardavimų didinimo ir klientų išlaikymo, tačiau egzistuoja prieigos prie duomenų ir jų kokybės problemos.
- Žmogiškųjų resursų vaidmenį (human in the loop). Visi ekspertai pabrėžė, kad DI sprendimas pradiniam etape neturėtų būti 100% autonomiškas - reikalinga specialistų priežiūra, kadangi egzistuoja informacijos spragos bei „haliucinacijų“ rizika.

6.1.2. Prekių katalogų duomenų analizė

Po interviu buvo atlikta aptartų įmonių e. prekybos svetainių duomenų analizė:

- Nustatyta, kad daugelio produktų aprašymai yra nepakankamai detalūs arba pateikti ne vienu formatu (nestandardizuoti laukai, trūkstamos specifikacijos).
- Paaiškėjo, kad ribotas kiekis struktūruotos informacijos gali lemti didesnę DI modelio „haliucinacijų“ tikimybę.
- Tapo akivaizdu, jog, net ir nesiimant visos duomenų kokybės gerinimo iniciatyvos, reikėtų suprojektuoti sprendimą, kuris galėtų panaudoti tik turimus duomenis, o trūkstamas vietas leistų įvertinti ir patikslinti žmogui („human in the loop“ principas).

6.1.3. Rezultatas

Šiame etape paaiškėjo, jog norint sėkmingai pritaikyti DI personalizacijos sprendimą, reikia spręsti ribotos produkto informacijos (4.2 skyrius) iššūkį ir numatyti integracijos su esamomis sistemomis scenarijus (4.3 skyrius).

6.2. Rinkos poreikių tyrimas ir konkurentų sprendimų analizė

Siekdami interpretuoti gautus duomenis ir interviu įžvalgas platesnėje rinkoje, vykdėme sistemingą mokslinės literatūros bei viešų šaltinių analizę. Tikslas - identifikuoti, kokie personalizacijos sprendimai jau egzistuoja, kokios yra pagrindinės tendencijos bei galimos nišos mūsų kuriamam maketui.

6.2.1. Mokslinės literatūros ir statistinių šaltinių analizė

Nagrinėdami viešas ataskaitas (pvz., [redstagfulfillment.com](https://www.redstagfulfillment.com), [globenewswire.com](https://www.globenewswire.com)), patvirtinome, kad bendra e. komercijos rinka auga itin sparčiai - 2023 m. pardavimų vertė siekė apie 6,5 trilijono USD, o prognozuojama, kad iki 2027 m. ji išaugs iki beveik 10 trilijonų USD.

Išsami rinkos analizė parodė, kad vartotojai tikisi greitai rasti aktualias prekes. Analizėje pastebėta, kad maždaug 12% vartotojų palieka svetainę (angl. *cart abandonment* ir bendras naršymo nutraukimas) dėl pernelyg didelio produktų pasirinkimo, neaiškios

navigacijos arba nepakankamo asmeninimo ([algolia.com](https://www.algolia.com)). Tai įrodė, kad esama paieškos sistema neatitinka šiuolaikinių klientų lūkesčių.

Kituose tyimuose ([globenewswire.com](https://www.globenewswire.com), [statista.com](https://www.statista.com)) pabrėžiama, kad net iki 68% vartotojų nori gauti asmeninių rekomendacijų internetu, o daugiau nei 40% pirkinių atsisakymų įvyksta dėl informacijos trūkumo arba per didelio pasirinkimo.

6.2.2. Konkurentų ir rinkoje egzistuojančių sprendimų apžvalga

Analizavome įvairius rinkoje siūlomus DI sprendimus ar paieškos papildinius (pvz., [Searchanise](https://www.searchanise.com), [Bloomreach](https://www.bloomreach.com), [Shopify AI](https://www.shopify.com) integracijos), kurie naudoja generatyvinį DI arba paprastesnius rekomendavimo algoritmus.

Daugumoje jų akcentuojamas personalizacijos pranašumas (tarkime, vartotojo naršymo istorijos pritaikymas, kad pirmiausia būtų rodomi jau anksčiau pirktų prekių atnaujinimai arba susijusios prekės).

Tačiau taip pat paaiškėjo, kad dauguma šių sprendimų susiduria su dviem esminėmis problemomis:

- Didelių pradinių investicijų poreikis (ypač vidutinėms ar mažesnėms įmonėms), apie kurį rašoma [searchanise.io](https://www.searchanise.io) tyimuose.
- Trūkstanti ar nestruktūruota duomenų bazė, dėl kurios DI negali tiksliai siesti produktų su vartotojų poreikiais (tą pabrėžia [thevirtualforge.com](https://www.thevirtualforge.com)).

6.2.3. Tendencijų ir pagrindinių iššūkių identifikavimas

Personalizacija tapo beveik neišvengiamu reikalavimu šiuolaikiniuose e. komercijos sprendimuose - tyrimai rodo, kad asmeniškai vartotojui pritaikyti pasiūlymai kelis kartus padidina konversijas, o didelės įmonės (pvz., „Amazon“) net ~35% savo pardavimų gauna iš rekomendacijų variklio ([evdelo.com](https://www.evdelo.com)).

Tačiau net didelės kompanijos susiduria su vadinamąja DI „haliucinacijų“ problema (4.1 skyrius) - pavyzdžiui, Perplexity AI atvejis parodė, kad modeliai gali supainioti skaitinius dydžius ar kitas svarbias produkto detales ([hypotenuse.ai](https://www.hypotenuse.ai)).

[Bloomreach](https://www.bloomreach.com) ir [McKinsey](https://www.mckinsey.com) duomenimis, nors DI diegimas gali padidinti pajamas daugiau nei 25% ar net iki 40%, būtina užtikrinti vartotojų duomenų privatumą ir nenutrūkstamą modelių treniravimą, antraip nuostoliai dėl netinkamų rezultatų ir vartotojų nepasitenkinimo gali atsverti naudą.

6.2.4. Rezultatai

- Atlikta analizė patvirtino, kad personalizacija yra būtina šiuolaikinei e. komercijai, siekiant mažinti vartotojų pasitraukimą ir didinti pardavimus. Tačiau ši sritis vis dar patiria didelių iššūkių, susijusių su duomenų kokybe ir DI modelių patikimumu.
- Iš konkurentų ar analogiškų sistemų peržvalgos aiškiai matyti, kad vien tik „išmanus“ algoritmas nepakanka, jei duomenys nėra tvarkingi ir struktūruoti. Būtent todėl mūsų tikslingas dėmesys „human in the loop“ metodikai, pradiniam testavimui vidinėje aplinkoje ir duomenų kokybės auditas pasirodo esąs ne tik novatoriškas, bet ir reikalingas norint sumažinti diegimo rizikas.

Taigi, rinkos poreikių tyrimas ir konkurentų analizė ne tik pateikė pagrįstus statistinius rodiklius (apie rinkos dydį, vartotojų elgseną, personalizacijos svarbą), bet ir parodytą patirtį, patvirtinančią, kad sprendžiami iššūkiai (ribota produktų informacija ir DI „haliucinacijos“) yra aktualūs visai rinkai. Surinktos įžvalgos suformavo pagrindą tolesniems technologiniams bandymams (6.3 skyrius) ir vartotojo sąsajos maketavimui (6.5 skyrius).

6.3. DI technologinių variantų analizė

Šiame etape, įvertinus pirminių interviu (6.1 skyrius) bei rinkos poreikių tyrimo (6.2 skyrius) išvadas, buvo sprendžiama, kaip geriausiai pritaikyti didelius kalbos modelius (LLM) personalizuotoms paieškos ir rekomendacijų sistemoms kurti. Konkrečiai svarstyti du esminiai modelių tipai:

- **„Komerčiniai“** (uždari) modeliai/paslaugos - pavyzdžiui, OpenAI, Amazon Lex, Google Vertex AI ir kiti panašūs sprendimai, kuriuose vartotojai gali naudotis modelių funkcionalumu per API, bet neturi priėjimo prie pačių modelio svorių ar vidinės architektūros.
- **Atviro kodo** (*open source*) LLM - pavyzdžiui, Llama, Falcon, Mistral, GPT-Neo, ir kt., kurių išeitiniai (angl. weights) failai yra prieinami viešai arba su tam tikromis licencinėmis sąlygomis, leidžiančiomis juos diegti vietinėse (*on-premise*) ar specializuotose infrastruktūrose ir pritaikyti savo reikmėms.

6.3.1. Privalumų ir trūkumų įvertinimas

	Komerčiniai modeliai/paslaugos (OpenAI, Amazon Lex ir kt.)	Atvirojo kodo LLM (Llama, Falcon, Mistral ir kt.)
Privalumai	<ul style="list-style-type: none"> • Reikia prižiūrėti mažiau infrastruktūros, naudojama kaip paslauga • Teikėjas nuolat atnaujina modelius • Greitesnis pateikimo rinkai laikas • Reikia mažiau darbuotojų specializacijos • Pigiau, jei mažai naudojasi (mokama už tokenus) 	<ul style="list-style-type: none"> • Visiška vidaus duomenų, naudojamų tiksliai derinimui ir (arba) RAG, kontrolė • Gali būti įdiegtas lokaliai, jei to reikalauja teisės aktai • Ekonomiškai efektyvesnis, kai apkrova yra didelė, o ilgalaikės išlaidos nuspėjamos
Trūkumai	<ul style="list-style-type: none"> • Gali būti brangu, jei daug naudojama • Gali būti atskleisti jautrūs duomenys - užklausos gali būti naudojamos būsimiems kitų modelių versijų mokymams. Išimtis taikoma įmonių klientams ir „Premium API“ naudotojams • Priklausomybė nuo tiekėjo 	<ul style="list-style-type: none"> • Galimai didesnės pradinės diegimo išlaidos • Reikalingi kvalifikuoti darbuotojai, kad būtų galima įdiegti ir prižiūrėti • Manoma, kad modelio kokybė yra prastesnė, palyginti su „OpenAI“

6.3.2. Eksperimentinis bandymas

Praktiniais sumetimais buvo išbandyti keli atvirojo kodo modeliai (tarp jų Llama ir Falcon variantai) bei keli komerciniai sprendimai (OpenAI ChatGPT, Anthropic Claude). Testavimo tikslas buvo:

- Išmatuoti tikslumą teikiant rekomendacijas ar atsakant į klausimus, turint ribotą produkto ir kliento informaciją.
- Įvertinti greitį (atsakymo sugeneravimo laiką).
- Nustatyti sąsajų (*integration*) paprastumą su išorinėmis sistemomis (pvz., ERP, PIM, Exacaster Customer360).

Apibendrinus eksperimentus, **paaiškėjo**, kad:

- Net geresni komerciniai modeliai linkę „**haliucinuoti**“, jeigu trūksta duomenų apie konkretų produktą ar nėra aiškių ribų, iš kokių šaltinių galima semtis žinių.
- Atvirojo kodo modeliai suteikė daugiau lankstumo, tačiau reikalavo didesnio diegimo, priežiūros ir tobulinimo laiko.
- „**Human in the loop**“ koncepcija tapo kritiškai svarbi, siekiant patikrinti generuojamų atsakymų tikslumą, užtikrinti kokybės kontrolę bei mokyti modelius iš realių korekcijų.

6.3.3. Sprendimas dėl maketo



6.3.3.1 pav. Sprendimo loginė schema

Pastebėjome, kad net pažangūs modeliai periodiškai generuoja neatitinkančią realybės informaciją (haliucinuoja) arba neteisingai interpretuoja trūkstamus duomenis. Todėl nusprendėme **integruoti „human in the loop“ praktiką**, kad specialistai galėtų patvirtinti arba koreguoti modelio generuojamas rekomendacijas prieš jas pateikiant galutiniams vartotojams.

Paaiškėjo taip pat, jog didelės apimties modeliai naudoja daugiau resursų, tad pirmame etape tikslinga šiuos sprendimus diegti ribotai - pavyzdžiui, pradėti nuo vidinių darbuotojų, kurie gali greičiau suteikti grįžtamąjį ryšį ir padėti tobulinti modelio logiką.

Duomenų privatumas ir atskirties nuo viešų šaltinių klausimai yra aktualūs, todėl sprendime bandomas „Agentic AI“ metodas, pagal kurį modelis turi ribotą prieigą tik prie tam tikrų specializuotų duomenų šaltinių (pvz., *Exacaster Customer360*), o ne prie visos internetinės bazės.

6.4. Antrieji kokybiniai interviu su kitais ekspertais

Paaiškėjo, kad per greitas DI įrankių atidavimas galutiniams e. parduotuvės lankytojams kelia daugiau rizikos (galimos haliucinacijos, nepatikimi rezultatai, reputacinė žala), nei atneša apčiuopiamos naudos. Todėl pasitvirtino hipotezė, jog optimalu DI sprendimą iš pradžių skirti vidiniams įmonės darbuotojams (parduotuvės konsultantams, klientų aptarnavimo specialistams), kurie galėtų tiesiogiai dirbti su DI, stebėti atsakymus ir iš karto

juos taisyti ar atsisakyti klaidingos rekomendacijos, taip suteikdami itin vertingą grįžtamąjį ryšį modelio tobulinimui.

Papildomai apklausėme tris kitus įmonių atstovus, siekdami gilesnio supratimo, kaip DI galėtų padėti vidiniam personalui. Šįkart kalbėjomės su:

- Customer Service vadovu (e-komercijos srityje),
- Telemarketing & Point of Sales padalinio atstovu (telekomunikacijų srityje),
- Customer Care skyriaus vadovu (logistikos srityje).

6.4.1. Rezultatai ir pagrindinės išvados

- *Customer Service* atstovas pabrėžė, kad ilgas informacijos rinkimo procesas dėl fragmentuotų sistemų žymiai prailgina užklausų sprendimo laiką ir padidina klaidų riziką.
- *Telekomunikacijų ir mažmeninės prekybos* atstovas išreiškė poreikį turėti lengvą ir vieningą produktų paieškos mechanizmą, kur DI galėtų užpildyti esamas informacijos spragas.
- Visi paminėjo, kad dažna problema - kai klientas susisiečia su klausimu dėl konkretaus pasiūlymo ar nori palyginti kelis produktus, tačiau informacija apie atskiras produktų savybes būna nepilna ar reikalauja papildomos paieškos išorinėse sistemose (pvz., Google).
- Bendras įžvalgų rinkinys nurodė, kad didesnė DI nauda yra pasiekama vidiniam naudojimui, nes patyrę darbuotojai veikia kaip papildomas kokybės kontrolės etapas, užtikrinantis tikslumą prieš pateikiant rekomendacijas klientams.
- Pagrindiniai sėkmės rodikliai, jų manymu, yra: *First Contact Resolution*, vidinis darbo efektyvumas (konsultacijų laiko sutrumpinimas), galimybė greitai integruoti paiešką su jau naudojama klientų valdymo sistema.

6.4.2. Poveikis projekto kryptčiai

Šių papildomų interviu dėka patikslinta projekto strategija:

1. **Trumpalaikė perspektyva:** DI sprendimas skirtas vidaus naudojimui (klientų aptarnavimo, telemarketingo, pardavimų komandoms), su integruota „human in the loop“ kontrole, kad būtų užtikrinta kokybė ir leidžiama modelį aktyviai tobulinti.
2. **Ilgalaikė perspektyva:** Po sėkmingo vidaus naudojimo sistemos bus papildomai optimizuojamos, tvarkomi duomenys (pvz., PIM/ERP integracija) ir tuomet galima bus išplėsti sprendimą į išorinius kanalus (galutiniams vartotojams).

6.5. Vartotojo sąsajos maketavimas ir testavimas

Remiantis ankstesnėmis išvadomis (tarp jų 6.4.2 punkto pabrėžta trumpalaikė perspektyva sprendimą naudoti tik įmonės viduje), nuspręsta sukurti ir ištestuoti pradinį sprendimo maketą, kuriame vartotojo sąsaja veikia pokalbių platformos (*chatbot*) pagrindu. Taip suteikiama galimybė įmonės darbuotojams (konsultantams, telemarketingo atstovams) natūralia kalba pateikti užklausas sistemai, kuri grąžina personalizuotas rekomendacijas.

6.5.1. Maketo kūrimo eiga

6.5.1.1. Žemo detalumo (low-fidelity) maketai

Pirmiausia buvo sukurti elementarūs, eskiziniai brėžiniai, kuriuose pavaizduota:

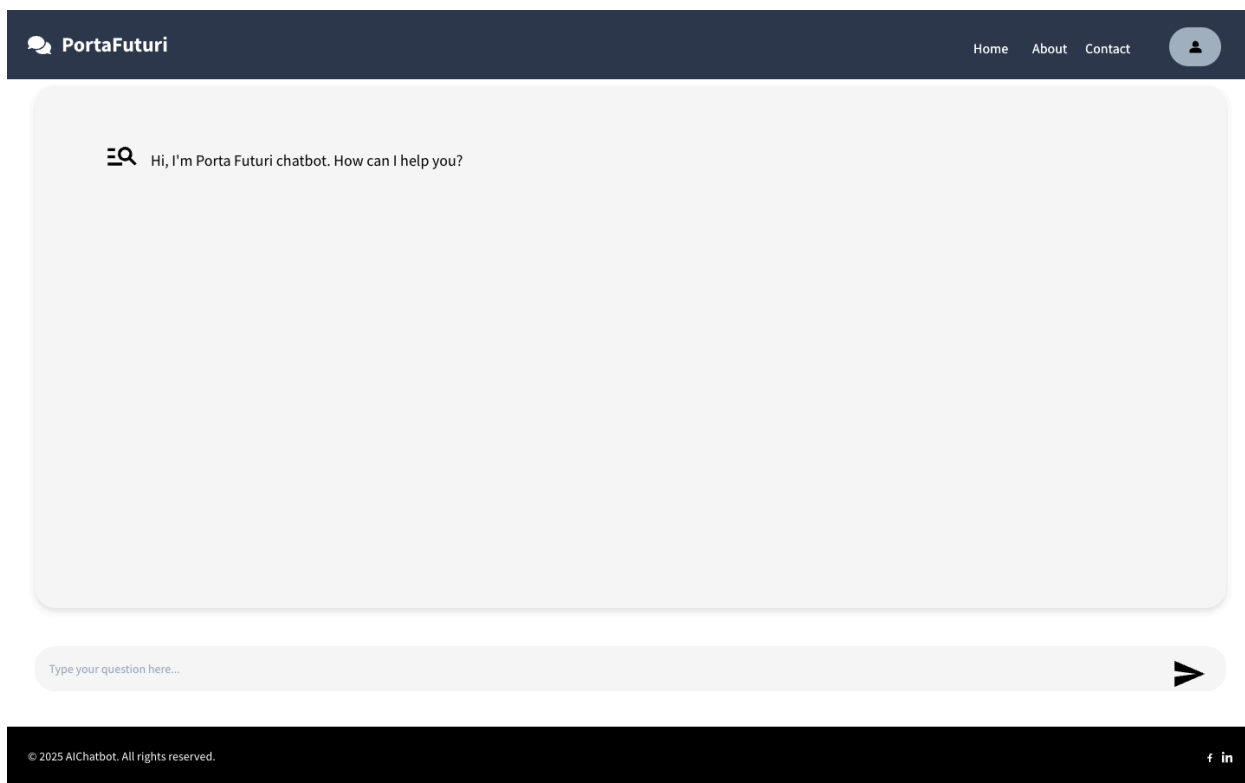
- Pokalbio langas (*chat window*), kuriame darbuotojas gali rašyti klausimus apie prekes ir/ar paslaugas.
- Paieškos rezultatai išdėstomi viename stulpelyje tiesiog su pavadinimu ir trumpu aprašymu.
- Minimalūs valdikliai (pvz., mygtukas „Patikslinti užklausą“ ar „Pasirinkti kitą pasiūlymą“), leidžiantys darbuotojui koreguoti užklausą arba atmesti netinkamą pasiūlymą.

Šie žemo detalumo *wireframe* variantai testuoti nedidelėje vidinių darbuotojų grupėje, kad būtų patikrintas bendras sąveikos su vartotojo sąsaja principas. Atsižvelgus į pastebėjimus (pvz., „reikėtų aiškiau atskirti kliento tekstą ir sistemos atsakymą“, „reikalingi papildomi paaiškinimai, kodėl tokia rekomendacija pateikiama“), *wireframe* buvo iteratyviai tobulinami.

6.5.1.2. Aukšto detalumo (high-fidelity) maketai

Atlikus pradinį eskizinių *wireframe* testų, buvo parengta detalesnė vartotojo sąsajos (*chat UI*) versija, kuri labiau priartinta prie galutinės, realiai naudojamos sistemos. Ji apima šiuos esminius elementus:

- **Pagrindinis langas**
Darbuotojas/konsultantas viršutinėje arba centrinėje dalyje mato pokalbio laukelį, į kurį galima natūralia kalba įvesti užklausą, pavyzdžiui:
„Ką rekomenduoti klientui 94569?“
Vieta, kur pateikiami tiek paties konsultanto, tiek DI atsakymai, stilistiškai atskiriama (pvz., naudojamos skirtingų spalvų kalbos „burbulai“).

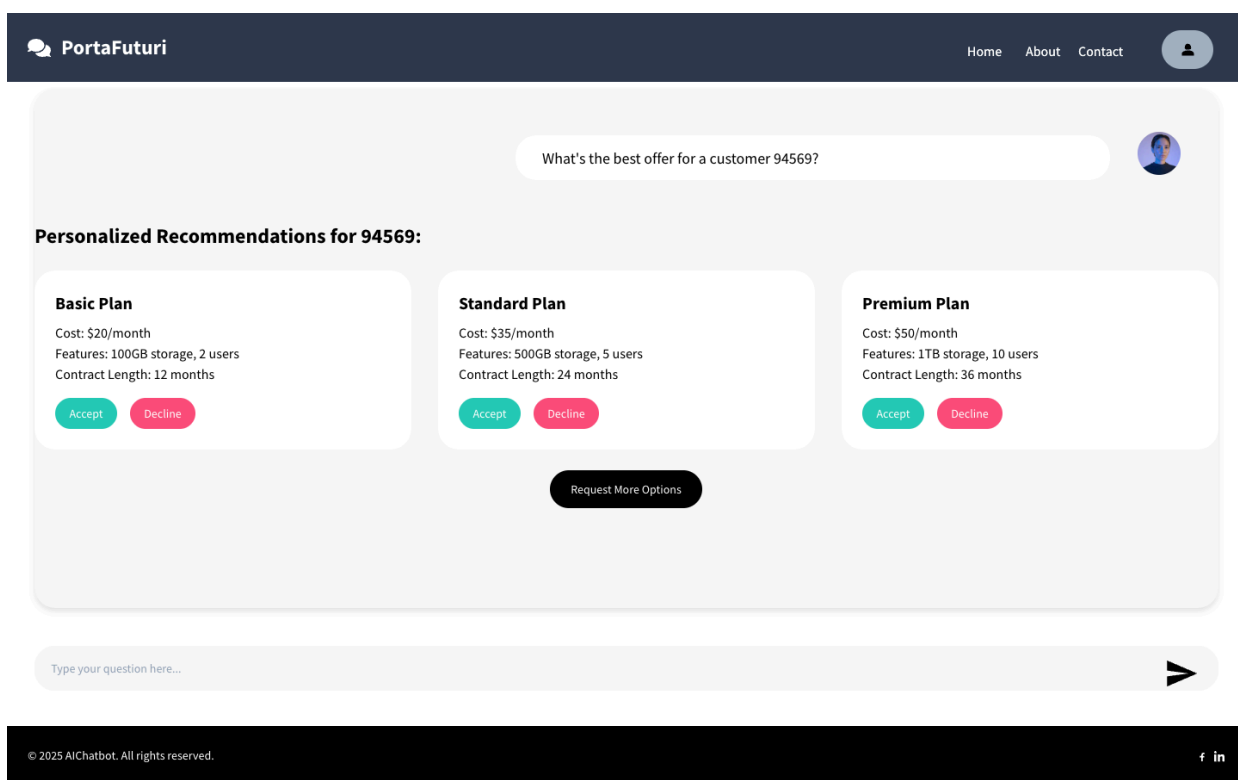


6.5.1.2.1 pav. Maketo pradinis langas

● Personalizuoti pasiūlymai

Pateikus užklausą, DI sistema fone parenka tinkamus pasiūlymus (planus, produktus) iš prijungtos duomenų bazės ir rodo galimų variantų sąrašą. Tai gali būti sąveikios žinutės, kur konsultantas gali:

- Perskaityti trumpą pasiūlymo aprašymą, kainą.
- „Perkelti“ pasiūlymą į atmintinę ar padaryti kitą veiksmą (pvz., spustelėti mygtuką „Daugiau informacijos“).



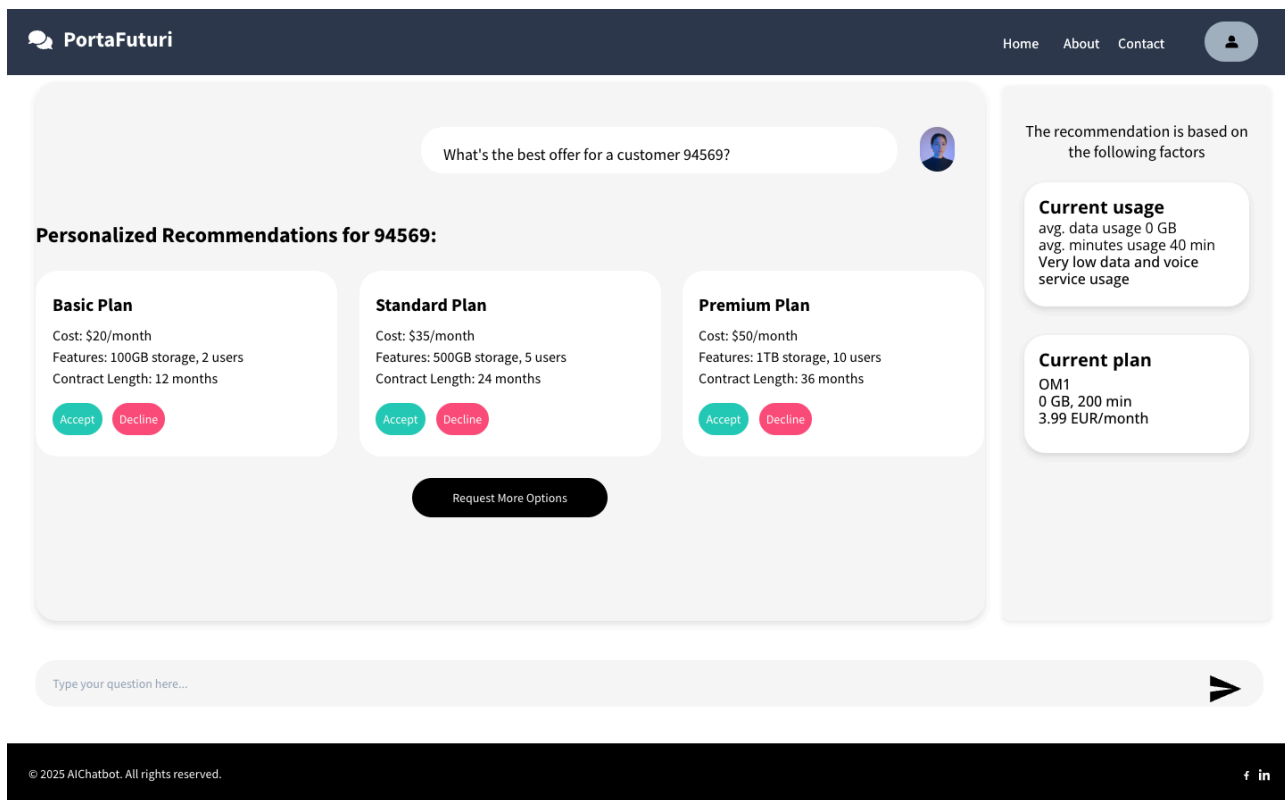
6.5.1.2.2 pav. Maketo langas, kuriame atvaizduojami personalizuoti pasiūlymai

- **Informacinė šoninė panelė (info panel)**

Kartu su pasiūlymais rodomi paaiškinimai, kodėl DI siūlo būtent tokią prekę/paslaugą, pavyzdžiui:

„Klientas šiuo metu turi planą X, tačiau dažnai viršija duomenų kiekį, todėl siūlome planą Y su daugiau gigabaitų...“

Informacinėje panelėje matyti pagrindinės kliento naudojimo tendencijos arba esama įranga (jei kalbama apie telefoną). Tokia skiltis padeda konsultantui greičiau suprasti sistemos logiką ir pagrįsti savo sprendimą klientui.

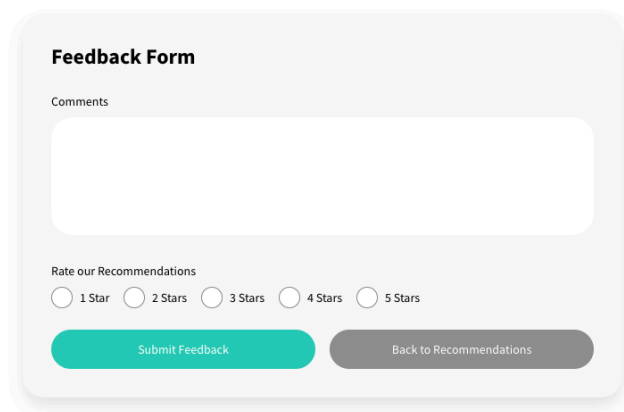


6.5.1.2.3 pav. Maketo langas, kuriame atvaizduojama informacinė šoninė panelė

- **Konsultanto atsiliepimo mechanizmas**

Norint užtikrinti „human in the loop“ principą, *high-fidelity* maketuose įdiegta galimybė greitai vertinti pasiūlymus:

- Mygtukas ar žymė „Patvirtinti pasiūlymą“ (jei konsultantas mano, kad DI atitiko kliento poreikį).
- „Atmesti ir pakoreguoti“, leidžiantis nurodyti priežastį (pvz., „kaina per didelė“, „klientas nemėgsta šio gamintojo“).
- Ši informacija vėliau pateikiama DI modeliui, kad būtų renkami duomenys tolesniam mokymui ar prompt tobulinimui.



6.5.1.2.3 pav. Maketo langas, kuriame konsultantas įveda atsiliepimą

Tokie aukšto detalumo maketai padėjo labiau atskleisti, ar darbuotojams pakanka vizualios informacijos apie produktus/paslaugas ir kokia forma jie norėtų veikti (pvz., reikalingos žinutės su piktogramomis, greitosios atsakymų parinktys ar pan.). Taip buvo suformuota ir vartotojo sąveikos (UX) logika, artimesnė realioms naudojimo situacijoms.

6.5.1.3. Human in the loop

Nors pokalbio sąsajos (*chat UI*) tikslas yra kuo labiau automatizuoti pasiūlymų teikimo procesą, atliekant maketo (*wireframe*) testavimą tapo akivaizdu, jog dalyvaujantis žmogus (konsultantas, telemarketingo darbuotojas) yra būtinas siekiant kokybiškos ir patikimos paslaugos. Tai ypač aktualu tuo atveju, kai:

- Klientas turi specifinių poreikių: yra situacijų, kai net pažangus modelis negalės „nuspėti“ kliento norų (pvz., klientas kategoriškai nesinaudoja mobiliais duomenimis), tad žmogus, žinodamas papildomas aplinkybes, gali atmesti netinkamą pasiūlymą.
- Reikalingi paaiškinimai: info panelė aiškina, kodėl pasiūlymas pateiktas, bet jeigu konsultantas vis tiek mano, kad tai nepakankamai motyvuota, žmogus gali perklausti ar patikslinti užklausą, taip kartu tobulindamas DI.

Į maketą „*human in the loop*“ sprendimas įsilieja per kelis elementus:

- **Greitas koregavimas:** po kiekvieno DI atsakymo konsultantas gali duoti grįžtamąjį ryšį, t.y. patvirtinti, atmesti, pakoreguoti.
- **Korekcijos priežastis:** naudotojas gali nurodyti, kodėl rekomendacija netinkama; ateityje ši priežastis padės pagerinti modelio logiką.

- **Nuolatinis tobulinimas:** kaupiant informaciją apie pasitaikančius netikslius ar klaidinančius atsakymus, tampa aišku, kuriose srityse reikalingi patobulinimai (arba duomenų bazėje, arba modelio prompt nurodymuose).

Todėl, net ir siekiant didelio DI automatizavimo, šiame projekte „human in the loop“ išlieka esmine kokybės užtikrinimo priemone - taip garantuojama, kad galutinis sprendimas bus geriau pritaikytas klientų poreikiams ir sumažės rizika dėl galimų modelio klaidų.

6.5.2. Maketo testavimas su vidiniais įmonės darbuotojais

6.5.2.1. Testinių scenarijų pavyzdžiai

Aptariant maketą (kai jau dalis dizaino elementų detalizuoti, bet dalis vis dar eskiziniai) - <https://app.uizard.io/p/dbe8f734>, tyrimo dalyviams buvo pateikiami šie pagrindiniai klausimai.

Maketo sritis	Klausimai
Pokalbio eiga ir bendravimo su sąsaja supratimas	<ul style="list-style-type: none"> • „Ar jums aišku, kaip pradėti pokalbį su sistema ir suformuluoti užklausą?“ • „Ar pakanka <i>chat</i> tipo žinučių sekos, ar norėtumėte aiškesnės žingsnių sekos?“
Pasiūlymų rodymas	<ul style="list-style-type: none"> • „Ar tinkamai matote, kokius pasiūlymus sistema rekomenduoja? Ar jie pateikiami priimtinu būdu?“ • „Ar pasiūlymų išdėstymas padeda greitai nuspręsti, kuris variantas tinkamiausias klientui?“ • „Ar pakanka pateikiamos kainos ir savybių informacijos?“
Info panelės paaiškinimai	<ul style="list-style-type: none"> • „Ar pakanka paaiškinimų, kodėl DI sistema siūlo būtent šią paslaugą ar prekę klientui?“ • „Kokios papildomos detalės (pvz., kliento planų istorija, dabartinė įranga) dar būtų naudingos priimant sprendimą?“ • „Ar aiškiai suvokiate, kur paspausti, kad matytumėte išsamesnę informaciją apie pasiūlymą?“
„Human in the loop“ veiksmas	<ul style="list-style-type: none"> • „Ar pastebite, kur galėtumėte atmesti ar redaguoti sistemos pasiūlymą? Ar tai darytumėte to pačio pokalbio metu, ar norėtumėte atskiro mygtuko?“

	<ul style="list-style-type: none"> • „Ar patogų pažymėti priežastį, kodėl atmetate pasiūlymą (pvz., per brangu, netinkamas gamintojas)?“ • „Ar numatyti patvirtinimo / atmetimo mygtukai atitinka jūsų darbo logiką?“
Elementų išdėstymas	<ul style="list-style-type: none"> • „Ar aišku, kam skirtas kiekvienas maketo blokas (chat langas, pasiūlymų blokas, šoninis info panel)?“ • „Kuriuos elementus norėtumėte perkelti, keisti dydį ar galbūt paslėpti?“ • „Ar užtenka vaizdinės hierarchijos, kad atskirtumėte sistemos atsakymus nuo savo paties pranešimų?“
Bendra vartotojo patirtis	<ul style="list-style-type: none"> • „Ar užtrukote ne per ilgai, kol supratote, kaip naudotis?“ • „Ar manote, kad sistemos atsakymų rodymas tinkamas? Ar norėtumėte, kad jie būtų pateikti išsamiau / glaustai?“ • „Kas labiausiai padėjo jums suprasti, ką daryti toliau šioje sąsajoje?“
Papildomi pasiūlymai ir komentarai	<ul style="list-style-type: none"> • „Kiek jums svarbu matyti krovimo / laukimo pranešimus, kol sistema apdoroja kliento duomenis?“ • „Ar norėtumėte kitokių filtrų, kategorijų, kurios padėtų greičiau suvaldyti pasiūlymą?“ • „Kokių papildomų funkcijų pasigendate (pvz., galimybės generuoti PDF pasiūlymą klientui, siųsti nuorodą e. paštu)?“

6.5.2.2. Surinkti rezultatai

Remiantis surinktais darbuotojų atsakymais bei stebėjimais sesijų metu, išryškėjo kelios bendros tendencijos:

- Vartotojai greitai suprato pokalbio eigą, nes analogiškos *chat* sąsajos jiems įprastos (pvz., bendraujant su klientais ir kolegomis per pranešimų platformas). Tačiau kilo klausimų dėl tikslios sekos, kai reikia patvirtinti / atmesti pasiūlymą – dalis dalyvių norėjo atskiro mygtuko, kitiems buvo aišku naudotis tuo pačiu pokalbio lauku.
- Pasiūlymų rodymas kortelių stiliumi sulaukė teigiamo įvertinimo, nes taip lengviau palyginti kelis variantus. Tačiau kai kurie dalyviai minėjo, kad norėtų, jog pasiūlymų

lentelę būtų galima išskleisti / suskleisti, ypač jei klientui įdomūs keli kriterijai. Dalis norėjo matyti tik vieną, pagrindinį pasiūlymą.

- *Info panel* - svarbi, bet norėtusi dar konkretesnių paaiškinimų, kaip sistema priėjo prie vieno ar kito rekomendacijos varianto. Tai rodo, kad norima detalios istorijos apie kliento naudojamus planus, ypač telekomunikacijose.
- Ne visiems vartotojams buvo intuityvu, kad jie gali atmesti pasiūlymą tiesiog pakomentavę, todėl ateityje rekomenduojama daryti akivaizdžius mygtukus „Netinka klientui“, „Nepatraukli kaina“ ar pan.
- Daliai dalyvių norėjosi, kad pokalbio langas ir pasiūlymų sąrašas nebūtų rodomi tame pačiame stulpelyje, nes tai gali užkrauti ekraną. Siūloma dalį informacijos perkelti į šoninę panelę ar iššokantį langą.
- Dauguma testuotojų gana greitai perprato maketo logiką (tiesa, reiktų atkreipti dėmesį, kad tai visi vidiniai darbuotojai, turintys bazinį supratimą apie projektą). Dalis pasiūlė aiškesnę atskirtį tarp konsultanto klausimų ir DI atsakymo burbuliukų (pvz., pakeisti spalvų kodus).
- Kol kas testuotojai žinojo, kad tai tik wireframe, todėl nesitikėjo visiško funkcionalumo. Visgi dalis minėjo, jog būtų naudinga matyti „krovimo“ indikatorių, kad suprastų, kada modelis apdoroja duomenis.
- Vieni dalyviai pageidavo galimybės parengti iš anksto suformuotus šabloninius atsakymus klientams. Kiti pasigedo greitų filtrų (pvz., pagal kainą, pagal duomenų kiekį), kad agentas net nedėtų neaktualių pasiūlymų.

6.5.2.3. Korekcijos

Remiantis šiais rezultatais, makete:

- **Patobulinta pasiūlymų blokų išdėstymo hierarchija**
Pvz., svarbiausias, labiausiai kliento poreikį atitinkantis planas matomas pirmiausia, o likusieji variantai rodomi tik vartotojui pageidaujant (pasirinkus „Rodyti daugiau“, „Rodyti pigesnę/prabangesnę variantą“). Taip sumažinamas informacijos perkrovimas ir konsultantui lengviau susikonsultuoti į geriausią pasiūlymą.
- **Išsamiau detalizuoti argumentai, pagrindžiantys personalizuotą rekomendaciją**
Kadangi dalyviai pageidavo geriau suprasti, kodėl sistema pasirinko vieną ar kitą pasiūlymą, maketo *Info Panel* skiltis išplėsta. Ten pateikiama, pavyzdžiui, „Šis planas geriau tinka klientui, nes jis vidutiniškai sunaudoja 12 GB duomenų per mėnesį, o dabartiniame plane gauna tik 10 GB“. Toks paaiškinimas suteikia konsultantui daugiau pasitikėjimo rekomendacijomis ir padeda argumentuotai bendrauti su klientu.
- **Optimizuotas atsakymo formatas, kad konsultantas galėtų jį tiesiogiai pateikti klientui**
Pastebėta, jog norint greitai aptarnauti klientą, svarbu, kad konsultantas nedėtų papildomų pastangų perrašant ar perfrazuojant sistemos atsakymą į kliento kalbą. Todėl rekomendacijos aprašymas pateikia natūraliai suformuluotą tekstą (pvz., „Jums labiausiai tiktų Planas X, nes...“), kurį konsultantas gali beveik pažodžiui

perskaityti ar nukopijuoti į komunikacijos kanalą. Taip taupomas laikas, mažėja rizika netiksliai perteikti informaciją.

Apibendrinant, šie patobulinimai leidžia konsultantui efektyviau panaudoti DI sprendimą, greičiau suprasti rekomendacijos logiką ir operatyviai pateikti aiškų atsakymą klientui, todėl gerėja ir kliento, ir paties konsultanto patirtis.

6.3. Maketui pritaikyta architektūra

Norint, kad kuriami *wireframe*, t.y. maketai, veiktų ne tik kaip eskiziniai paveikslai, bet ir leistų realiu laiku sąveikauti su įmonės duomenų bazėmis, inžinieriai parengė atitinkamą architektūrinę koncepciją:

1. Kliento pusė (*Front-end*)

Chat sąsaja, atkartojanti *wireframe* dizaino elementus. Vartotojo įvestas tekstas (užklausa) siunčiamas į vidinę API, o gautas DI atsakymas pateikiamas tame pačiame vartotojo sąsajos lange.

2. Tarpinis sluoksnis (*Middle layer*)

Apdoroja vartotojo žinutę, kreipiasi į DI agentą, o prieš pateikdamas galutinį atsakymą, atlieka tam tikrus patikrinimus (pvz., ar duomenys aktualūs, ar modelis nepasiūlė neegzistuojančios prekės).

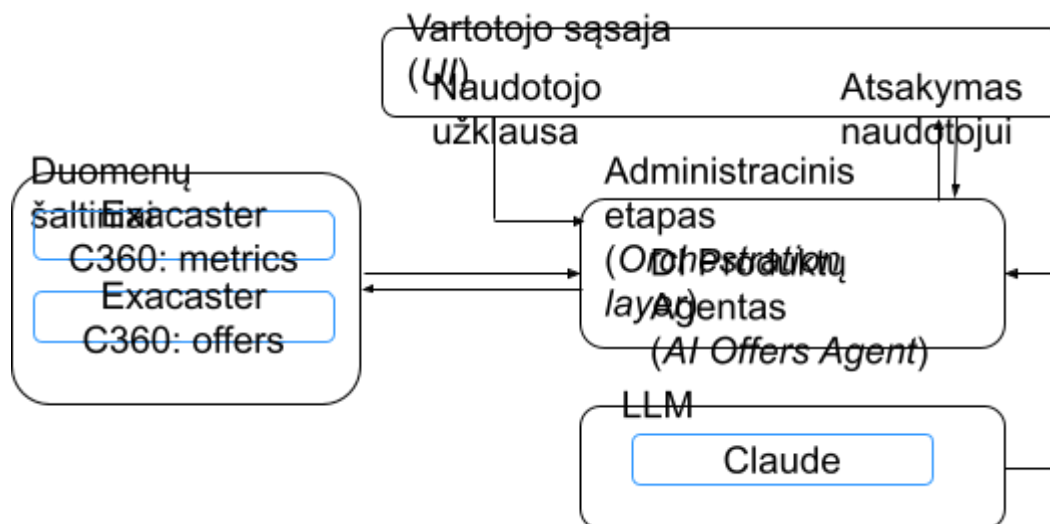
3. DI agentas

Pats DI logikos branduolys, integruotas su ribotais šaltiniais (planų, prekių, kainų API). Agentas formuoja personalizuotą atsakymą remdamasis klientų elgsenos duomenimis, kuriuos laiko Exacaster Customer360, bei kitais prijungtais įrankiais.

4. “Human in the loop”

Užtikrinama galimybė, kad bet kuriuo metu konsultantas galėtų atmesti DI pasiūlymą ir sugrįžti su patikslinta užklausa. Sistemoje registruojama priežastis, o surinkti duomenys naudojami modeliui ar užklausų (*prompt*) šablonams tobulinti.

Toks architektūrinis sprendimas tiesiogiai paremtas *wireframe* maketo reikalavimais, kadangi visa sąveika vyksta per pokalbio langą, stengiamasi, kad duomenų srautai ir DI atsakymo logika būtų aiškūs ir greitai reaguotų į vartotojo įvestą tekstą.



6.3.1 pav. Architektūros koncepcija, maketo įgyvendinimui

Bendra struktūra užtikrina, kad LLM modelis nepasitelktų jokių išorinių šaltinių (kurie galėtų lemti „haliucinacijas“) ir būtų koncentruotas į realius, sistemose kaupiamus duomenis.

6.4. Tyrimo rezultatų išvados ir tolesni veiksmai

Tyrimo eigos rezultatai užtikrina, kad 2 skyriuje apibrėžtos užduotys - nuo rinkos analizės iki maketo kūrimo ir testavimo - buvo įgyvendintos, o 4 skyriuje įvardinti techniniai iššūkiai sprendžiami, taikant duomenų kokybės vertinimo priemones, *human in the loop* kontrolę bei nuosekliai planuojant integracijas su vidinėmis sistemomis.

Tyrimo eigos metu gauti rezultatai:

- **Pirminiai interviu ir duomenų analizė** padėjo aiškiai identifikuoti praktines spragas: ribota produktų informacija, integracijų trukdžiai, DI „haliucinacijų“ rizika.
- **Rinkos poreikių tyrimas ir konkurentų sprendimų analizė** atskleidė, kad personalizacija yra strateginis prioritetas, tačiau daugeliui sprendimų trūksta aiškaus būdo, kaip valdyti netikslius (ribotus) duomenis ir užtikrinti patikimumą.
- **DI technologinių variantų analizė** įrodė, jog net pažangūs modeliai gali klysti, todėl ankstyvoji fazė turi būti orientuota į vidinių darbuotojų poreikius, - taip sukurama galimybė testuoti ir tobulinti DI be didelio reputacinio pavojaus išorinėje rinkoje.
- **Maketo testavimas** su skirtingo detalumo (*low* ir *high fidelity*) versijomis parodė, jog *chat* principu paremtas sprendimas konsultantams yra patogus, tačiau būtina aiškiai atskirti rekomendacijas ir pateikti jų pagrindimą. Įvykdytos korekcijos (pasiūlymų hierarchijos patobulinimai, išsamesni argumentai ir pritaikytas atsakymo formatas) leidžia sumažinti darbuotojo pastangas bei užtikrinti operatyvesnį bendravimą su klientu. Toks rezultatas priartina personalizuotą DI pasiūlymų modelį prie realios kasdienės naudojimo aplinkos.

Tolesniuose etapuose planuojama:

- A. Sukurti naudotojo dialogo (*chat*) sąsają.

-
- B. Atlikti testavimą su integruotais duomenų šaltiniais ir vartotojo sąsajos aplinka.
 - C. Atlikti *benchmarking* tyrimą, kai lyginami DI sprendimai su panašiomis rinkoje esančiomis sistemomis, siekiant įvertinti mūsų sprendimo konkurencingumą.
 - D. Įvertinti galimybes dalį sprendimų pateikti galutiniams vartotojams (po to, kai bus surinkta pakankamai patvirtinto grįžtamojo ryšio).

Taip užtikrinama, kad DI sprendimas ne tik efektyviai padės vidiniams darbuotojams, bet ilgainiui bus tinkamas ir platesniam komerciniam naudojimui, atitinkant visus privatumo, patikimumo ir konkurencingumo reikalavimus.

● 7. Pasiiekti rezultatai

Šiame skyriuje pristatomi pagrindiniai projekto metu pasiekti rezultatai, atskleidžiantys tiek technologinio sprendimo vertę, tiek tolimesnes jo tobulinimo galimybes.

7.1. Sukurtas vartotojo sąsajos maketas

Projekto metu buvo sukurtas interaktyvus hibridinis (*low* ir *high fidelity*) pokalbio tipo vartotojo sąsajos maketas, leidžiantis įmonės darbuotojams testuoti DI pagrįstą rekomendacijų sistemą. Maketas integruoja DI agento atsakymus, personalizuotą informacijos pateikimą bei galimybę darbuotojui palikti grįžtamąjį ryšį pagal „human in the loop“ principą.

7.2. Patvirtintas „human in the loop“ naudingumas

Kita svarbi projekto išvada - „human in the loop“ (žmogiškojo faktoriaus) metodika pasirodė itin veiksminga:

- **Žmogiškojo ir DI sąveikos sinergija**

Pažymėtina, kad toks DI sprendimas kol kas tinkamesnis vidiniam naudojimui, nes vidiniai darbuotojai gali įvertinti, ar gautas atsakymas tikslus. Jie veikia tarsi žmogiškoji kontrolė, užtikrinanti, jog klientui persiunčiama informacija yra teisinga. Taigi, net jei modelis pateikia dalinai netikslų atsakymą, patyręs specialistas gali jį pakoreguoti prieš perduodant klientui. Šis žmonių ir DI derinys pasiteisina, panašiai kaip kitose pramonės šakose, kur derinamos žmonių ir DI stiprybės kuriant vertę klientams (bcq.com). Tai ypač svarbu, jei duomenų kokybė ne visuomet ideali arba DI modelis dar turi polinkį „haliucinuoti“.
- **Grįžtamasis ryšys**

Kiekvieną kartą, kai darbuotojas pastebi netikslumą (pavyzdžiui, pasenusią kainodarą ar kliento poreikiui netinkančią prekę) ir jį ištaiso, DI modelis gauna naudingos informacijos tolesniam tobulinimui (panašiai kaip kituose mašininio mokymosi sprendimuose, kai žmogus pakomentuoja modelio klaidas).

7.3. Gauti pirmieji testų rodikliai

Maketo testų metu užfiksuoti keli svarbūs kiekybiniai ir kokybiniai rodikliai:

- **Duomenų kokybės įtaka**

Viena svarbiausių išvadų - DI personalizacijos ir rekomendacijų kokybę riboja ne tiek modelio algoritminės galimybės, kiek **duomenų turinio kokybė**. Projekto metu identifikuota, kad srityse, kur duomenys nepilni, DI veikimas prastas - modelis arba visai nesugeba sugeneruoti konkretaus atsakymo, arba duoda miglotą rekomendaciją. Šis atradimas atitinka plačiai žinomą tendenciją: *DI sprendimų kokybė tiesiogiai priklauso nuo duomenų kokybės* (dqlabs.ai). Jei duomenys netikslūs ar neišsamūs, AI negali „atspėti“ teisingai.

- **Haliucinacijų mažinimo priemonės**

Remiantis pirminiais bandymais, efektyviausios strategijos, sumažinančios klaidinančių (haliucinuotų) atsakymų skaičių, yra:

- naujausių DI modelio versijų naudojimas - naujesnės versijos paprastai užtikrina pagerintus nuspėjimo algoritmus ir tiksliau apdoroja kontekstą.
- agentinis metodas („Agentic approach“) - suteikiant modeliui ribotą prieigą prie patvirtintų duomenų šaltinių (įrankių), gerokai mažėja tikimybė, kad DI prisigalvos nepatvirtintos informacijos.
- „Human in the loop“ - įtraukus ekspertinę žmogaus kontrolę ir grįžtamąjį ryšį, dar labiau sumažėja galutinių klaidų (ypač pradiniam etape, kai duomenų bazėje dar esama trūkumų).

- **Vartotojo sąsajos (chat UI) efektyvumas**

Chat sąsajos testavimo rezultatai parodė, kad konsultantai greitai perpranta maketo logiką ir geba ja naudotis. Pastebėta, kad svarbiausi sąveikos komponentai - aiškus rekomendacijos pagrindimas, galimybė atmesti pasiūlymą ir operatyvus atsakymo formatas, kurį galima persiųsti klientui - ženkliai padidintų vidinį pasitenkinimą sistema bei konsultacijos greitį.

7.4. Sukurtas tolesnis plėtros planas

Remiantis gautais rezultatais, identifikuotais tobulinimo aspektais ir 6.6 poskyryje išdėstyta tyrimo eiga, suformuotas aiškus veiksmų planas:

- **Duomenų šaltinių prijungimas**

Pradiniame etape sprendimas remsis Exacaster Customer360 platformos duomenimis, kurie pateikia pagrindinius kliento elgsenos duomenis (pvz., naudojamų planų istoriją, pirkimo įpročius), produktų ir paslaugų sąrašą. Tolesniuose žingsniuose planuojama prijungti kitas išorines sistemas, taip užtikrinant tikslesnius ir išsamesnius duomenis DI agento sprendimų priėmimui.

- **Įvairesnės integracijos**

Diegiant papildomus duomenų šaltinius (pvz., ERP, PIM, realiu laiku atnaujinamos kainodaros lentelės), planuojama užtikrinti nuoseklų duomenų kokybės augimą ir efektyvų sinchronizavimą, kad sistemos naudotojui būtų pateikiama tik naujausia ir tiksli informacija.

- **Dialogo (chat) sąsajos kūrimas**

Numatyta sukurti specialų pokalbio (chat) modulį, leidžiantį naudotojams bendrauti su DI agentu natūralia kalba patogesnėje vartotojo sąsajoje. Tai suteiks lankstesnę galimybę klausti ne tik apie konkrečius produktus, bet ir apie jų palyginimą, pritaikymą tam tikroms situacijoms.

Tokia sąveika supaprastintų „human in the loop“ procesą, t.y. darbuotojas ar testuotojas galėtų gyvai matyti agento atsakymą ir tiesiogiai jį koreguoti, taip tobulindamas sistemos supratimą.

- **Testavimo tąsa su išplėstais duomenimis ir atnaujinta UI**

Augant duomenų šaltinių spektrui, planuojama dar kartą išplėsti testus, kad būtų įtraukti nauji produktai, klientų istorijos, skirtingi naudojimo scenarijai. Bei kuriant naują chat modulį, bus vykdomi pakartotiniai bandymai, siekiant įvertinti, ar DI agentas išlieka spartus ir tikslus didėjant apkrovai bei duomenų apimčiai.

- **Konkurencingumo (benchmarking) tyrimas**

Planuojama lyginti sukurta maketą su rinkoje jau egzistuojančiais DI sprendimais (pvz., kitų e. komercijos paieškos ir rekomendacijų įrankių), siekiant objektyviai nustatyti, kokie yra mūsų maketo pranašumai ar trūkumai funkcionalumo, tikslumo ir greičio prasme. Gauti rezultatai padės identifikuoti sritis, kurias būtina toliau tobulinti (pvz., paieškos greitaveiką, duomenų interpretavimą ar pateikiamų rekomendacijų kokybę) ir pasisemti patirties iš kitų rinkos dalyvių.

- **Galimybių vertinimas dėl diegimo galutiniam vartotojams**

Pagrindinis rodiklis šio žingsnio sėkmei - patvirtinto grįžtamojo ryšio (iš vidinių darbuotojų, pilotinių vartotojų) kiekis ir kokybė. Jei testavimas rodo, kad atsakymų tikslumas bei sparta pakankami, sprendimą galima skalauti didesniai rinkos segmentui.

Taigi, šis plėtros planas apima tiek artimiausius techninius veiksmus (duomenų šaltinių integraciją, dialogo sąsajos kūrimą, tolimesnį testavimą), tiek strateginius žingsnius (benchmarking vertinimus, galutinio B2C diegimo svarstymą). Tokiu būdu nuosekliai pereinama nuo vidinio maketo visaverčio komercinio sprendimo, pritaikyto įvairiems e. komercijos scenarijams, link.

7.5. Verslo naudą patvirtinančios įžvalgos

Vidinis maketo testavimas patvirtino, kad integruota personalizuota paieškos sistema gali padidinti pirkimų konversijos rodiklius bei pagerinti klientų lojalumą, nes siūlomos rekomendacijos yra labiau pritaikytos konkrečiam vartotojo poreikiams.

Surinkti interviu duomenys rodo, kad sprendimas gali būti išplėstas ir pritaikytas kitose įmonės veiklos srityse (pvz., telemarketinge, klientų aptarnavime) - tai lems vieningos žinių valdymo platformos kūrimą ir tolesnį verslo augimą.

• 8. Veiklos vykdymo metu atliktų tyrimų eigos pakeitimai

Projekto vykdymo metu, atsižvelgiant į tarpinius rezultatus ir išmoktas pamokas, buvo padaryti keli tyrimų plano pakeitimai.

Apskritai, tyrimų eigoje nuolat taikytas Agile požiūris - atsižvelgiant į tarpinius rezultatus, greitai priimti sprendimai pakeisti arba papildyti tyrimo metodus. Šios adaptacijos leido optimaliai paskirstyti išteklius: pavyzdžiui, užuot gaišus laiką ties perfekcionistiniu modelio tobulinimu, dalis dėmesio nukreipta į išsamius rinkos tyrimus, kas buvo kritiška galutiniam sprendimo aktualumui.

8.1. Modelio pasirinkimo korekcija

Iš pradžių numatyta plačiau testuoti *atvirojo kodo* (open-source) LLM ir turėti visą infrastruktūrą įmonės serveriuose, tačiau paaiškėjo, kad per ribotą laiko tarpą greičiau bus pasiekiamas veikiantis modelis, taikant komercinę (uždarą) DI paslaugą. Taip buvo užtikrintas spartesnis startas ir koncentruotasi į maketo validavimą, užuot sprendus kompleksinius atvirojo kodo diegimo klausimus.

8.2. Metodologijos pritaikymas esant duomenų trūkumui

Pirminiame plane buvo numatyta koncentruotis į paties DI modelio (LLM) tobulinimą ir jo integracijos procesus. Tačiau jau ankstyvose bandymų fazėse paaiškėjo, kad pagrindinė kliūtis – ne tiek algoritmo apribojimai, kiek produktų duomenų kokybės trūkumas (nepilni aprašymai, nesuderinti atributai).

Atsižvelgiant į šią praktinę įžvalgą, tyrimo eiga buvo pakoreguota:

- Vietoje vieno atskiro dėmesio LLM parametrų ir prompt tobulinimui, nuspręsta į papildomą kokybinių interviu tyrimą įtraukti daugiau e. komercijos atstovų, duomenų valdymo specialistų bei potencialių projekto naudotojų.
- Šie kontekstiniai pokalbiai padėjo išsamiau suprasti, kodėl dalis produktų duomenų iš viso nėra suvedama, kurios produktų sritys labiausiai nukenčia dėl informacijos trūkumo ir kokios galimos alternatyvios išeitys.
- Gautos rekomendacijos leido greitai koreguoti prioritetus: didesnę dėmesį skirti vartotojo sąsajai, kuri (1) skirta vidiniams darbuotojams, (2) leidžia „human in the loop“ koreguoti ne tik galutinį DI atsakymą, bet ir pačius produktų duomenis.

Šis žingsnis pasirodė itin reikšmingas, nes patvirtino, jog DI sprendimų sėkmė pirmiausia priklauso nuo turinio (duomenų) kokybės - net patys pažangiausi generatyviniai modeliai negalės sukurti reikšmingų personalizuotų rekomendacijų, jei neturės patikimo ir nuolat atnaujinamo duomenų šaltinio.

8.3. Skirtingų modelio versijų palyginimas

Tirtas ne tik vienas pasirinkto LLM variantas, bet ir keletas to paties modelio versijų veikimas (pvz., Anthropic Claude 3.5 ir 3.7). Tokiu būdu buvo nustatyta, kad naujesnės versijos pasižymi spartesniu atsakymu ir mažesniu „haliucinacijų“ skaičiumi, todėl tyrime pradėtos taikyti naujesnės modelio versijos.

8.4. Testavimo apimtys korekcija

Iš pradžių buvo planuota vykdyti plačius vartotojų testus, įtraukiant išorinius klientus. Tačiau, įvertinus galimas rizikas ir konfidencialumo reikalavimus, šio plano atsisakyta. Nuspręsta apsiriboti vidiniu maketo testavimu, kuriame dalyvauja tik įmonės konsultantai bei kiti darbuotojai. Toks sprendimas sumažino tiesiogiai iš galutinių pirkėjų gautų atsiliepimų kiekį, tačiau leido saugiau ir patikimiau įvertinti sistemą ankstyvoje fazėje. Ateityje, kai modelio patikimumas bus didesnis, svarstoma grįžti prie galimybės įtraukti lojalius klientus į pilotinį testavimą - ši iniciatyva numatyta kitame projekto etape.

- **9. Kita su projekto įgyvendinimu susijusi informacija (ekranvaizdžiai, struktūrinės schemos ir pan.)**

Žemiau pateikiamos papildomos detalės ir iliustracijos, kurios padeda suprasti, kaip veikia maketas ir kaip jis buvo testuojamas.

9.1. Gerosios praktikos analizė ir analogų įžvalgos

Projekto metu buvo tikslingai remtasi ne tik akademiniiais šaltiniais ar konkurentų produktais, bet ir sėkmingais verslo atvejais, kuriuose integruotos DI sistemos derinamos su žmogaus kontrole („human in the loop“). Vienas iš pagrindinių pavyzdžių - „A1 Telekom Austria“ bendradarbiavimas su BCG (Boston Consulting Group) ([bcg.com](https://www.bcg.com)).

BCG rekomenduoja fokusuotis į tuos atvejus, kurie duos didžiausią investicijos grąžą (*ROI - return on investment*) - tą mes ir padarėme. Sudarėme matricą, kurioje viena ašis - poveikis verslui (didesnės pajamos, geresnė patirtis, kaštų taupymas), kita - įgyvendinimo sudėtingumas. Personalizuota paieška-agentų asistentas atsidūrė „aukšto poveikio, vidutinio sudėtingumo“ zonoje - tai reiškė, kad tai turėtų būti prioritetas naudojimo atvejis (ir mes jį realizavome pirma).

Kiti identifikuoti atvejai: personalizuotas turinio kūrimas (pvz., automatinis produktų aprašymų generavimas pagal vartotojo segmentą), išmanusis kryžminis (*cross-sell*) pardavimas (DI siūlo klientui komplektuoti susijusius produktus) ([bcg.com](https://www.bcg.com)). Šiems atvejams BCG analizėje priskyrėme antrinį prioritetą - juos planuojame nagrinėti kituose projektuose, kai pagrindinė paieškos funkcija bus įdiegta. Toks struktūruotas požiūris padėjo pamatyti *platesnį inovacijos vaizdą*: personalizuotos rekomendacijos nėra vieno įrankio diegimas, o visa naujų galimybių paletė, kurią reikia diegti nuosekliai, pradedant nuo labiausiai pasiteisinančių.

9.2. Sistemos architektūros vizualizacijos ir techninė konfigūracija

Architektūros koncepcija paremta agentinių DI metodu, kuriame LLM modelis neturi prieigos prie viso interneto, o dirba tik su ribotais, specialiai apibrėžtais duomenų šaltiniais. Tai leidžia:

- užtikrinti duomenų saugumą (atitikti BDAR);
- kontroliuoti žinių bazę (kad DI nepriimtų neteisingų prielaidų);
- efektyviai surinkti grįžtamąjį ryšį tobulinimui.

Vizualizacija pateikiama 6.3.1 pav.

9.3. Naudoti įrankiai ir technologijos

Komponentas	Aprašymas
LLM (Claude, GPT, Llama)	Naudoti tiek komerciniai, tiek open source modeliai – tikslumui palyginti
Jupyter Notebook (taip pat žinoma kaip IPython Notebook)	Sistemos architektūros modeliavimui
Wireframe įrankis: Uizard	UI/UX maketų generavimui ir testavimui su realiais naudotojais.

9.4. Projekto įgyvendinimo dinamika ir Agile adaptacijos

Tyrimo eigoje aktyviai taikytas Agile metodas:

- iteratyviai testuotas maketas,
- iteratyviai analizuota maketui įgyvendinti efektyviausia architektūra,
- prioritetai buvo keičiami pagal grįžtamąjį ryšį,
- sprendimai koreguoti realiu laiku pagal verslo poreikius (žr. 8 skyrių).

Tai leido ne tik laiku įgyvendinti pirminį maketą, bet ir suformuoti tvirtą pagrindą tolesnei plėtrai bei testavimui.